Internet Appendix to "Stimulating Housing Markets"

DAVID BERGER, NICHOLAS TURNER, and ERIC ZWICK*

This Internet Appendix contains material to aid replication and clarify the data build process. We also present supplementary analysis to demonstrate robustness of the paper's results and additional description of how we calculate aggregate program effects.

* Citation format: Berger, David, Nicholas Turner, and Eric Zwick, Internet Appendix to "Stimulating Housing Markets," *Journal of Finance*, Doi: 10.1111/jofi.12847. Please note: Wiley-Blackwell is not responsible for the content and functionality of any additional information provided by the authors. Any queries (other than missing material) should be directed to the authors of the article.

I. Data Build and Discussion

The analysis combines a large number of proprietary and public-use data sources. In this appendix, we describe each source in detail, describe variable construction, and walk step-by-step through sample selection.

A. Data Sources and Sets

1. *Tax records, IRS/OTA*: These data are anonymized individual-level data collected by the IRS for the purposes of administering the tax collection process. They are made available through collaboration with the Office of Tax Analysis in the U.S. Treasury Department and the IRS Division of Research, Analysis, and Statistics. The data include information about the age, earnings, marital status, number of dependents, and tax filing ZIP code reported on the income tax return.

   We compile the following items:

   (a) ZIP-5 level cross sections of FTHC claims from Form 5405. These include claims for versions 2 and 3 of the FTHC program (i.e., the grant program) and also claims for the Long-Time Homebuyer Credit.

   (b) First-time homebuyer and tax filer counts from individual tax returns and information returns for the years 1998, 1999, and 2000 to measure program exposure, and through 2013 to measure the age of first-time homebuyers in nonpolicy years.

   (c) Tax filer counts for 2007 from individual tax returns.

   Tax credit claims cover 18,073 ZIPs, the homebuyer counts from 2000 cover 24,923 ZIPs, and the 2007 records cover 35,647 ZIPs.

2. *DataQuick deed records*: We clean and merge data retrieved from DataQuick's assessor file, which contains information on individual properties used to assess property taxes, and DataQuick's recorder file, which tracks ownership changes and loans secured by properties.

   We begin with records from 2004 to 2013, up to the company's acquisition by CoreLogic. Assessor data cover 1,819 counties that account for 91.8% of the U.S. population. While not all counties tracked by DataQuick provide recorder data, 942 counties do and these match to 88% of all deeds tracked in the assessor data.

   Once the two files are joined, we produce our "canonical" list of transactions by applying the following filters:

   (a) Include only resales and new construction (types R and S in SR_TRAN_TYPE) that are arm's length[1];

[1] To our knowledge, the arm's length flag in DataQuick is the output of a "model" that classifies whether resales are genuine arm's length transactions. The model automatically excludes refinances and intermediate documents in a distress sale process. It attempts to exclude related-party transactions made at nonmarket prices, for example, because of divorce or bequest.

(b) Remove transactions between institutional buyers, such as two developers or a developer with a GSE;

(c) Remove middle-man sales where a property buyer sells the property on the same day;

(d) Keep only the transaction with the highest transfer value if there are duplicates with the same property, transaction date, buyer, and seller, with the intent of removing incomplete duplicate records.

We use the filtered, merged data to create three output data sets:

(a) *DataQuick transaction counts*: We aggregate the transactions in the recorder data to the (geographic unit)×(month)×(transaction type)×(distress type) level, where the geographic unit is a ZIP or county. For each level, we count the number of transactions.

This initial data set is divided into six distinct data sets for use in different parts of the analysis:

i. all sales (transaction types R and S, distress types ∈ NULL, A, I, S and O),
ii. nondistress resales (transaction type R, null distress indicator)[2]
iii. new housing/subdivision sales (transaction type S),
iv. purchases at foreclosure auctions (distress type A),
v. DataQuick-inferred short sales (distress type I),
vi. REO liquidations (distress type S).

A variation of the above aggregation splits counts to the (geographic unit)×(month) ×(transaction type)×(distress type)×(bedroom) level, where bedroom information comes from the assessor file. We mark each transaction as having missing bedroom information, one to three bedrooms, or four or more bedrooms. Geographies with more than 5% of months lacking transactions on properties with bedroom data, as well as geographies in states with consistently poor bedroom data collection, are dropped.

A second variation requires processing individual records further before aggregation:

i. Transactions on recently built housing are marked by checking the SA_YR_BLT column in the assessor file.
ii. Transactions on properties owned by developers or by government-sponsored enterprises are marked by running regular expression searches on the SR_BUYER column in the recorder file. The regular expressions for builders, developers, and GSEs are:

[2]Distress type O transactions appear to be classified as distress by DataQuick, but show no other sign of being a distress sale, such as an unusually low sale price or institutional seller. We exclude them by default, but confirm the robustness of our results by including them as nondistress transactions. These transactions make up at most 2% of the sample in policy-period months.

```
builder_re_list =
  ('HOME|BUILDER| BUILD|BLDR| HM(\s|S\s)| CONST |REAL ESTATE|',
   'PULTE|RYLAND|NVR|DR HORTON|CENTEX|LENNAR|MERITAGE|',
   'STANDARD PACIFIC')
developer_re_list =
  ('LLC|CORP|COMPANY| INC$|INTL| LAND |PROPERT(Y|IES)|TRUST|',
   'INVEST(OR|MENT)|LP')
gse_re_list =
  ('FNMA|FHLM|HUD-HOUSING|FEDERAL (H|NATIONAL)|FANNIE|',
   'SECRETARY OF (H|VET)|VETERANS AD')
```

iii. Transactions on properties considered recently in distress are marked by searching over a property's transaction history and checking if the property was distressed less than two years prior to the current transaction.

The exclusion of type L sales from the resales and the new housing data sets is due to DataQuick classifying properties with a later delinquent mortgage as type L, even if the sale associated with that mortgage was a nondistress sale. The label therefore cannot be used to infer whether the original transaction was distressed, not distressed, or otherwise unusual. However, these type L sales account for many transactions in the pre-crisis period, up to a majority in some places in California. For these reasons, they are included when constructing the price and loans data.

(b) *DataQuick price data*: For each geographic unit-by-month unit, we compute the group's median price based on the SR_VAL_TRANSFER column (for nondistressed properties). We exclude transactions for which price data are missing, about 10% of all transactions, from the computation. We code as missing any units for which fewer than five transactions are available. These price data serve to complement the CoreLogic price indices when the CoreLogic data are not available.

(c) *DataQuick loans data*: DataQuick counts mortgages attached to the paperwork on a closed sale as loans related to that transaction, and allows up to three liens connected to one transfer. However, these rows miss loans taken out by the buyer from different banks or subsequent to the purchase, which may include second liens and "piggyback loans" used to cover the cost of the down payment.

We connect second liens, which are categorized as refinances in DataQuick, to transactions by sorting data on SR_PROPERTY_ID and SR_DATE_TRANSFER, and then linking all refinances between property transactions with the date of the preceding transaction. Loan values on refinances taking place less than 100 days after the preceding transaction are added to the loan value on that transaction.

3. *CoreLogic prices*: CoreLogic Home Price Index (HPI) data from the national to the ZIP level were made available through the Initiative on Global Markets at Chicago Booth. Unlike the DataQuick price data, which records nominal values, the HPI is a variant of the Case-Shiller index measuring price changes in repeatedly transacted properties. The structure of the data is a balanced panel, available for 7,169 ZIP codes and 1,267 counties.

4. *FHFA prices*: FHFA Home Price Index (HPI) data from the national to the ZIP level. These data are public use. FHFA's price indices are available at the yearly level for the largest set of ZIPs in our sample and are based on repeat sales. Bogin, Doerner, and Larson (2016) describe the construction and data source of these price indices.

5. *Covariate data*: We construct a covariate data set from the 2010 American Community Survey, the 2000 Census, IRS public-use files at the ZIP level, and Equifax ZIP-level aggregates. The ACS data contain five-year averages (2006 to 2010) of demographic indicators at the ZIP level.

   From the Census we draw the fraction of census blocks classified as urban. From the ACS we draw population in 2007 and compute the average unemployment rate, the average of ZIP-level median age, the average of median rent, and the average fraction below the poverty line between 2006 and 2010. From the IRS we draw average gross income in 2005.

   Using aggregated ZIP-level counts from Equifax, we follow Mian and Sufi (2009) and define the subprime share as the fraction of consumers in a ZIP code in 1996 with credit scores below 660.

   We construct three control variables based on publicly available data disclosed under the Home Mortgage Disclosure Act (HMDA). For each measure, we first pull HMDA data on all originated, home-purchase loans for one- to four-family homes that are owner-occupied. These observations are then allocated from the census-tract level to the five-digit ZIP level using HUD-USPS crosswalk files.

   (a) The first variable estimates exposure to the Home Affordable Refinance Program (HARP) using the proportion of loan originations in a ZIP code that were purchased by Fannie Mae or Freddie Mac. For each origination, we create an indicator variable that equals one if the loan was purchased by Fannie Mae or Freddie Mac. The ZIP-level measure of HARP exposure is then based on the average share of loan originations purchased by Fannie Mae or Freddie Mac in each year from 2004 to 2007, weighted by the total number of originations in the ZIP in each year.

   (b) The next variable estimates the subprime share in each ZIP code based on the HMDA rate-spread variable. For each origination, we create an indicator variable that equals one if the rate-spread variable is greater than three percentage points. The ZIP-level measure of subprime exposure is then based on the average share of loan originations where the rate-spread variable is greater than three in each year from 2004 to 2007, weighted by the total number of originations in the ZIP in each year.

   (c) Finally, we estimate exposure to the expansion of FHA loan limits in 2008. First, we estimate expanded FHA loan limits in each ZIP based on the publicly available FHA loan limits from 2015. For counties that were not listed as high-cost by GSEs in 2009, the estimates are based on the change in the national floor from 2007 to 2008. For the smaller number of high-cost counties, we create an estimated mapping based on the loan limits for one-unit homes in 2015. Our final measure is then a count of the home loan originations from 2004 to 2007 that were above the estimated 2007 loan limits, but less than or equal to the estimated 2008 loan limits over the total number of home originations in that ZIP over the same time period.

We construct a measure of exposure to the Home Affordable Modification Program (HAMP) using data on mortgage modifications made available on the Treasury Department's HAMP website. This data set contains information on mortgage modifications at the Metropolitan Statistical Area (MSA) level. We allocate MSA modifications to ZIP codes proportionately based on a ZIP code's 2007 MSA share of owner-occupied homes and scale the ZIP count by the number of mortgagees in that ZIP code. We include all modifications from September 2009 to June 2013 as noted by the Modification Effective Date variable.

B. Sample Selection

We define the main analysis sample beginning with the nondistress resales data set described above. To ensure estimates are not biased by changes in geographical coverage, only ZIPs or counties with more than 90% of their transaction time series complete from 2006 onwards are included. This will tend to exclude very small ZIPs that have many months during which there are no transactions. All other data sets are filtered through an inner join that restricts the analysis sample to the same set of ZIPs.

Table IA.II details the creation of the nondistress resales data set.

II. Supplementary Analysis

A. Nonreversal Evidence

A natural question following from our result on the nonreversal of home sales in the post-policy period is whether this phenomenon is general or whether some market segments do indeed experience short-run reversal. To explore this idea, we divide ZIPs into groups based on local housing market and housing stock characteristics and estimate heterogeneous effects. The results are in Table IA.VII and discussed briefly in the main text in Section IV.A. We run long difference regressions that compare average monthly sales for the pre-period versus the post-period for the full sample and various subsamples.

We consider three dimensions of heterogeneity. First, we compare the bottom three and top three deciles in median house prices during 2008 as proxies for the income and liquidity needs of first-time buyers. Second, we compare the bottom three and top three deciles in the starter home share of sales in sample years prior to the program. This separation proxies for the "move up" buyers who tend to buy larger homes as their families expand. Third, we compare the bottom three and top three deciles in the average pre-policy holding time for first-time buyers, identified in the Federal Reserve Bank of New York Consumer Credit Panel (CCP), constructed from Equifax. This separation proxies for "move up" buyers through ex post realizations among buyers in different ZIPs.

Column (1) confirms the nonreversal in the full sample. Columns (2) and (3) suggest that there is more reversal in the high-price areas, though recall that the bulk of the overall response occurs in low- and medium-price areas. Columns (4) and (5) reveal little difference across low- and high-starter-share ZIPs, although the samples are small due to incomplete information on bedrooms across the full sample. Columns (6) and (7) show that areas with long average duration among first-time buyers show less reversal than areas with short average duration. The table includes statistical tests of equality of coefficients across groups, which indicate that we are not able to detect meaningful differences along these heterogeneity dimensions.

Overall, we interpret the results above as suggestive evidence that short-term reversal is more likely when induced buyers are higher income, not liquidity constrained, and participate in markets where buyers are more mobile. At the aggregate level, the relative share of such buyers versus longer-term buyers determines the extent and speed of reversal. The shift in the overall age distribution and the stronger quantity effects in lower price markets suggest a large number of longer-term, possibly constrained buyers were induced into the market, which supports our preferred interpretation of the nonreversal finding.

B. Macro-Level Price Evidence

Our micro-level estimates provide strong cross-sectional evidence that the FTHC increased house prices. While these results are informative about the effectiveness of the FTHC policy, their value in facilitating a welfare analysis is incomplete because the research design removes macro variation from the estimates. As a result, there is no necessary connection between the micro-level results and the effect of the program on aggregate house prices: it is possible for strong micro-level and weak macro-level effects to coincide. Thus, despite obvious limitations, examining aggregate house price data is useful for gaining insight into whether the FTHC succeeded at stabilizing the aggregate housing market.

Figure IA.6, Panel A plots the monthly CoreLogic national house price index (thick gray line) and three separate price indexes that we build from DataQuick transaction data for the period 2004 to 2013. The CoreLogic index is based on repeat-sales pairs and incorporates a three-month moving average to smooth the aggregate data. The raw repeat index (dotted blue line) follows published methodology from CoreLogic and S&P Case-Shiller for including sales pairs and estimating price indices. The moving average repeat index (thin red line) applies a three-month moving average to the raw repeat index. The raw mean price index (dashed black line) includes all transactions, winsorized at the 1% level to reduce the influence of outliers.

At first glance, there appears to be mixed evidence that the FTHC affected aggregate house prices. All four series show a precipitous decline in prices starting in 2007, which slows dramatically in the first quarter of 2009, consistent with a positive policy effect. The CoreLogic index shows two humps in price levels that appear to lag the credit's expiration dates. However, there is little evidence in any of the three repeat-sale price indices that prices appreciated considerably during the policy period. In contrast, the raw aggregate price index, which includes every transaction, suggests significantly greater price appreciation during the policy period and persistent effects after the policy expires. What explains the difference?

Relative to raw prices, the repeat-sales indices make several adjustments in estimation and sample selection, the effects of which are illustrated by our replication. These adjustments include applying a three-month lagged moving average smoother to the estimated series and down-weighting transaction pairs with higher predicted variance of price changes. Our moving average replication matches CoreLogic closely, but the raw repeat-sales index shows bumps in prices that more closely coincide with the spikes in aggregate transactions. In addition, the repeat-sales index excludes transactions for which the reference property sells only once during the estimation period, and the most commonly cited repeat-sales indices exclude multifamily units and condominiums (including our CoreLogic series and replication).

Taken together, these adjustments tend to obscure the effect of the program on aggregate prices, either by smoothing sharp changes in the time series or by excluding transactions that are more likely to be bought by first-time homebuyers or to see significant price increases during this time. Figure IA.6, Panels B and C show that only using repeat transactions to construct the price index is a significant restriction. This restriction excludes just under half of total arms-length transactions on average, with a bias toward recently built homes. Importantly, these excluded transactions are not missing at random over time: during the policy period, the share of excluded transactions closely mirrors the time-series pattern of aggregate sales shown in Figure 2.

Many of the housing transactions that were induced by the policy are excluded from the price index, potentially biasing our estimated price effects downward. The rows of Table IA.VIII confirm these suspicions. Estimated price effects in the raw data are approximately twice the size of our baseline specifications, though with larger standard errors. Price effects using a hedonic index based on these data (first row) demonstrate robust price impacts.[3] Estimates using raw price growth (second row) are noisier due to the small set of transactions that drive ZIP-level means. We conclude that the price effects of the program, especially as visible in aggregate data, are likely obscured by the methodology of standard price indices.

C. *Recently Built Homes and Homes in Distress*

We exploit the richness of the DataQuick transaction data, which record the names of buyers and sellers as well as categories for distressed sales, to explore how likely it was that the FTHC induced reallocation of houses from low- to high-utility users. For each property, DataQuick's county assessor data provide detailed information on the characteristics of the transacted homes, including price, size, age, and the number of bedrooms and bathrooms. Linked to each property are transaction data that track changes in deed. The data identify whether a transaction is a short sale, a financial institution-owned sale (REO), a foreclosure auction, or an exchange of deed on a foreclosed home. DataQuick's proprietary model classifies whether the transaction is made between related parties or at arm's length.

It is important to distinguish sellers who are homeowners from those left holding assets they were unable to sell. A number of negative externalities are associated with the

[3]The hedonic index corrects raw price data for quality differences as a function of deciles of lot size, ventiles of square footage, ventiles of age, and indicators for the number of bedrooms and bathrooms (rounded to the nearest integer).

latter case. Empty houses decay more rapidly and can be subject to vandalism or host to other crimes. Foreclosure spillovers associated with forced sales of distressed homes can depress housing values for neighbors and, through subsequent reappraisals, amplify barriers to refinancing.

We divide the total transactions for our main analysis sample during the policy period into categories based on the likelihood that the sellers were not first-best users of the homes they were selling. We investigate the following categories:[4]

1. *Recently Built*: Includes homes built between 2005 and 2010.
2. *Short Sale*: Includes homes categorized by DataQuick as short sales (i.e., sales involving principal forgiveness by lenders).
3. *Foreclosure/Real Estate Owned (REO)*: Includes homes categorized by DataQuick as being sold from a financial institution's portfolio of homes or through a foreclosure auction.
4. *Developer Seller*: Includes homes for which the seller is either a home builder or other kind of company, based on the seller's name.
5. *Government Sponsored Enterprise (GSE) Seller*: Includes homes for which the seller is a federal entity—Fannie Mae, Freddie Mac, Ginnie Mae, the Federal Home Loan Banks, or the Veterans Administration.

During the policy period, there were 4.3 million total transactions in our sample. Approximately 739,000 of those were homes that had been built in 2005 or later. This compares to 396,000 homes classified as new construction by DataQuick. To the extent that our new construction marker is too restrictive, these sales indicate that the program may have allowed builders to sell homes from their recent inventories. Consistent with this view, approximately 1.13 million homes, or 24% of all home sales, were sold by developers or builders. Thus, nearly two-thirds of homes sold by developers were not new construction.

Recent construction does not contribute to output or employment at the time of sale. However, the importance of recent construction in aggregate sales during this period highlights two macroeconomic issues created by investment overhangs. First, an overhang of previously built assets reduces investment today while the economy redeploys excess inventory. Second, while GDP correctly measures the delivery of new homes during the construction period, it does not correctly measure the initiation of a stream of consumption services if those assets are subsequently left vacant. The facts suggest that programs like the FTHC can work by accelerating redeployment and initiating use of idle assets.

Distressed sales and sales from financial institution portfolios were also important during this time. Within our sample, there were approximately 561,000 short sales and 843,000 foreclosure or REO sales, including 235 thousand sales from the government entities' portfolios of repossessed homes. While not mutually exclusive from the recent construction and developer sales above, these figures again suggest that many of the transacted homes did not involve transfers from one homeowner to another, but instead enabled transitions of underutilized assets to more productive use.

[4]Internet Appendix Section I provides more details on how we categorize transactions, including regular expressions used to identify builders, developers, and the GSEs.

D. *Constrained Buyers and Mortgage Finance*

A large literature documents the importance of down payment constraints in housing markets. Stein (1995) shows that modeling down payment constraints is crucial for matching many empirical features of the housing market. Using the Panel Study of Income Dynamics (PSID), Engelhardt (1996) shows that young households reduce consumption in years in which they buy a home and increase consumption back to long-run levels in subsequent years. This suggests that the down payment constraint is binding for many young households. Survey evidence confirms this fact. Fuster and Zafar (2016) administer a survey on the role of down payment constraints on household willingness and ability to buy housing. They find that a reduction in down payment would have a much larger effect on household behavior than a decline in mortgage rate. This result reflects the difficulty many households face in saving for the typical 20% down payment, especially in areas with high home prices.[5]

The FTHC program coincided with an expansion by the Federal Housing Administration (FHA) of its first-time homebuyer mortgage guarantee program. This program enables mortgage loans of up to 96.5% of purchase price for eligible buyers. Given the low down payment requirements, first-time homebuyers make up a significant portion of new originations supported by the FHA, as the government-sponsored enterprises Fannie Mae and Freddie Mac typically require larger down payments. According to the Department of Housing and Urban Development (HUD), FHA supported 781,000 first-time homebuyers during 2009 and 882,000 during 2010, or approximately 56% of the first-time buyer market during these years.[6]

These low down payment loans are not costless: a lower upfront payment trades for higher subsequent interest payments plus required mortgage insurance premiums. A simple calculation highlights the trade-off. Consider three different mortgage contracts for a house that costs $200,000: (1) a conventional 30-year fixed rate mortgage requiring a 20% down payment, (2) a 96.5% loan-to-value (LTV) FHA loan where the household pays off the Upfront Mortgage Insurance (UMI), 1.75% of price, within the down payment, (3) a 96.5% LTV FHA loan that shifts the UMI into the principal. The FHA loan also includes a 0.55% mortgage insurance premium. Assume that the interest rate is 4.8%, the average conventional mortgage rate from November 2009. Last, assume that buyers expect to carry the mortgage for five years, the median duration among first-lien, first-time mortgages computed for 2009 and 2010 vintage loans in the LPS McDash loan servicing data set.[7]

---

[5] A report by builderonline.com finds that residents making the median income in a state have to save nearly eight years on average to put 10% down for a median price home (http://www.builderonline.com/money/how-long-will-buyers-save-up-for-the-down-payment-of-their-dream-home_o, Accessed: 2016-06-15). Similarly, a survey by Trulia.com finds that a full 47% of surveyed renters would consider buying if they had enough savings for the down payment (http://www.trulia.com/blog/trends/trulia-american-dream-survey/, Accessed: 2016-06-15).

[6] See Figure 6 in HUD's "Annual Report to Congress" (Department of Housing and Urban Development, 2011).

[7] For the sake of simplicity, we abstract from mortgage prepayment and exclude the tax benefits of

Under these assumptions and with no discounting, the first mortgage would cost \$50,200 over five years. Notice that to receive this contract the buyer makes a \$40,000 down payment at origination. The second mortgage would cost \$65,600 over five years but the buyer would only make a \$7,000 + \$3,500 down payment upfront. Finally, the third mortgage including UMI would cost about \$66,800 over five years and the buyer would make a \$7,000 down payment. In 2009, the average interest rate on FHA loans was 1.4% higher than a conventional mortgage, which increases the cost of the FHA loans to \$75,700 and \$77,100, respectively. Thus, the FHA mortgages are considerably more expensive than a conventional mortgage. That so many households chose an FHA mortgage despite the higher future cost suggests that down payment constraints were highly relevant for these households during the sample period.[8]

Part of the FHA expansion came from increased loan limits through the FHA Forward program initiated as part of the Economic Stimulus Act of 2008. This contemporaneous policy raises the concern that we incorrectly attribute effects of the FHA policy change to the FTHC program. However, Figure IA.4 shows that areas targeted by this policy were higher-price areas not particularly exposed to the FTHC. As noted above, controlling for ZIP-level exposure to the FHA policy change does not alter our estimates. Thus, the more relevant part of the FHA expansion for our purposes came through loans that qualified under the pre-crisis regime. This expansion occurred as the private market for second liens to reduce down payments contracted.[9]

A final piece of evidence of binding constraints comes from the patterns of FTHC claims via amended tax returns. By amending a prior return, buyers could accelerate the credit's payment to the time of purchase, instead of waiting until filing the next year's return. For most buyers, waiting until next year's return meant receiving the credit in May or June of 2010, potentially a year after purchase. The data reveal that the majority of buyers strongly preferred to receive their credit immediately, as 53% of claims came via amended returns. In addition, this share is decreasing in age. Among buyers aged 30 or younger, the amended share is 57%. In contrast, only 48% of buyers aged 40 and older claimed the credit via amended return. These patterns suggest that immediate liquidity demands, which are typically higher for young households, were important to FTHC claimants.

---

interest payments and insurance premiums. Duration data include voluntary prepayments only. FHA loans have slightly longer durations than conventional loans, as buyers have relatively limited refinancing options.

[8]Conversations with economists at HUD confirm that borrowers were permitted to apply the credit toward the 3.5% down payment.

[9]See HUD Mortgagee Letter 2009-07 and Goodman, Seidman, and Zhu (2014) for more details on the FHA loan limit expansion. A distinct mortgage assistance program called FHA Secure, created in August 2007 and expanded in July 2008, is also unlikely to bias our results, as this program only enrolled a few thousand borrowers.

E. Vacant Homes and Household Formation

To further assess the likelihood of beneficial reallocation, we complement our exploration of low-utility sellers from Section C with information from the de-identified tax returns of FTHC claimants. Unfortunately, the tax data do not record information about the people or entities from whom FTHC claimants bought their homes. Nor can we use aggregate data to assess this question since the effect of housing demand on aggregate vacancies is ambiguous.[10] However, it is possible to use information about mailing addresses to ask two related questions about claimant transitions into homeownership. The first question is whether the home occupied by the claimant at the time of purchase was occupied in recent years or instead a vacant home. The second question is whether the transition induced household formation in the sense that claimants move from a multiple-occupancy household to a single-occupancy household.

We attempt to measure vacancy and household formation using mailing address information reported on an individual's tax return.[11] To capture changes in vacancy status, we examine whether the new address associated with the FTHC purchase had been occupied two years prior to the purchase. To measure changes in household formation, we count the number of tax returns filed from a particular address and compare it to the number of tax returns filed at the FTHC claimant's address two years prior to the purchase. In both cases, we choose the period two years prior because the new address may be assigned to the prior year's tax return if a tax filer amended the prior return to claim the credit. We focus on claims made for purchases in 2009 to separate first-time homebuyers from long-time homebuyers.

From the FTHC claims in 2009, we find that 42% move into an address that had no filers in 2007 and 33% transition into a single-tax-filer address from living in a multiple-filer address in 2007.[12] We also compute these statistics for first-time homebuyers in the nonpolicy years between 2002 and 2013. The data suggest that FTHC claimants are not more likely to move into vacant homes, but are more likely to form new households relative to first-time buyers in other years (33.1% transition to single family in 2009 relative to 30.5% in other years). In years prior to the crisis, high vacancy rates may instead reflect new construction. We address this issue by dividing first-time homebuyer addresses based on whether the address appears in previous years. Using this method, on average 13.9% of the addresses between 2002 and 2007 can be classified as new construction. In 2009, only 7.9% of addresses can be classified this way, lending further support to the notion that

[10]While an increase in the demand for housing lowers vacancies, all things equal, an increase in expectations of future housing demand could lead to an increase in vacancies if more current homeowners choose to list their home.

[11]Specifically, we use the mailable point information encoded in the 12-digit ZIP code. We restrict analysis to valid ZIP-12s, that is, ZIP-12s for which the last seven digits are not all zeros.

[12]The household formation statistic restricts the sample to those filers for whom we have a valid previous address. As a validation check of the vacancy data, we confirm that the vacancy share of FTHC claims at the ZIP level is strongly correlated with the share of home sales that are foreclosures or short sales.

FTHC buyers often bought recently vacated homes.

Taken altogether, to the extent that the program accelerated reallocation of underutilized assets, this reallocation came primarily through increasing the level of home purchases during a time when the supply of vacant homes was abnormally high. Nevertheless, the data do rule out the possibility that FTHC purchases merely resulted in people "swapping" houses. In addition, the data suggest imputed owner-equivalent rental income as another indirect GDP effect of the program.

F. *Default Rates for Policy-Period Buyers*

Given the high origination LTV ratios of policy-period homebuyers and the literature suggesting that such LTVs can lead to distress, it is important to ask what happened to these buyers in the post-period. We use the DataQuick transaction data to shed light on this question. DataQuick records track a distressed property as it goes through each step of the default process, as early as a short sale and as late as the REO disposal following foreclosure. We use chronological changes in ownership classified as distress sales by DataQuick to identify homebuyers who later defaulted on their loan. Specifically, we follow policy-period cohorts of buyers and compare them to cohorts both before and after the policy. We restrict analysis to buyers with FHA-insured mortgages to focus on those with high potential default risk.

Figure 10 plots cumulative distress cohorts for purchases made during the policy period and compares these to cohorts based on 2006, 2007, and 2008 sales as well as cohorts based on 2011 sales. Our data allow us to compare cohorts for at least 36 months from the month of purchase. Both the 2009 and 2010 policy cohorts show no difference in default rate relative to the 2011 post-policy cohort. At 36 months out, each of these groups shows distress transition rates of approximately 10 per thousand purchases. Furthermore, all three of these groups display considerably lower rates of transition into distressed sales than the pre-policy groups. Thus, the data do not indicate that the FTHC program drew unusually risky buyers into the market, despite the very high LTVs at which these buyers entered. In this sense, the reallocation of homes appears to have been stable.

III. Detail on Aggregate Effect Calculations

A. *Aggregate Effect on Home Sales*

This section discusses three methods for estimating the aggregate quantity effect of the FTHC. The results are summarized in Table VII.

Method 1: Cross-sectional approach. Following Mian and Sufi (2012), we estimate the total number of sales caused by the program, exploiting only differences in cross-sectional exposure and using the group receiving the smallest shock as a counterfactual. We choose the bottom 1% of ZIPs as the counterfactual group, which corresponds to the bottom row of the heatmap in Figure 5, Panel A. We then compute the effect of the policy for other groups relative to this group. By construction, any time-series effect of the policy shown by the bottom group is set to zero and removed from the effect computed for other groups.

Standardized exposure is 0.85 for the bottom group and increases to 7.58 for the highest group. Thus, for each exposure group g, the aggregate number of sales induced by the program is

$$\Delta Sales_g = 17 \times \beta \times (e_g - 0.85) \times s_{g,2007}, \tag{IA1}$$

where β is the coefficient from Table III for the 17-month policy period, and e_g and $s_{g,2007}$ are weighted-average program exposure and monthly sales in 2007 for group g, respectively, where the weights are average monthly sales for the ZIPs in each group. We aggregate $\Delta Sales_g$ across all groups to provide an estimate of the aggregate effect within the DataQuick sample.

We estimate that the FTHC increased existing home sales by 169,000 units within sample during the policy period, or 8.1% of all sales during this period. In 2007, our sample covers approximately 41% of the national existing home sales market. Extrapolating our estimates to the national market yields an estimated increase of approximately 412,000 units during the policy period. There were 2.74 million claims of the FTHC during this time with a total outlay of \$20.3B.[13] Thus, under the assumptions that the lowest-exposure group is a plausible control group and that our sample is representative of the national market, 15% of claims were made for induced sales, as opposed to sales that would have happened in the absence of the policy.

Note that this is a lower-bound estimate if the lowest exposure group also responds to the program. A less conservative approach, which aggregates estimates relative to a zero-exposure baseline, yields an estimate of 233,000 units within sample, or 11.2% of all sales during the policy period. The extrapolated aggregate estimate in this case is 568,000 units and implies that 21 percent of claims were made for induced sales. The implied direct fiscal costs based on these two approaches are approximately \$49K and \$36K per induced sale, respectively.[14] Our preferred aggregate estimate equals 490,000 units, the average of these alternative approaches.

Method 2: Regression kink (RK) approach. We can also obtain an estimate of the total size of the FTHC program using the micro-level estimates from our RK design. Our preferred specification implies that the full FTHC credit increased the first-time homebuyer propensity by 0.76% among all tax filers and 3.2% among eligible tax filers. We use these individual elasticities to generate an estimate of the aggregate effect of the policy.

[13] This figure excludes approximately 550,000 claims for the Long-Time Homebuyer Credit.

[14] Our primary specification is a reduced-form regression of sales on exposure, and so delivers an intention-to-treat (ITT) estimate. Two-stage estimates with total claims per filing unit as the endogenous first-stage variable yield a first-stage coefficient, $\beta_{1S} = 0.568(0.048)$, and a policy-period second-stage coefficient, $\beta_{IV} = 5.26(1.72)$. This coefficient implies that a one-standard-deviation change in claims (0.018) leads to an increase of scaled sales of 0.095; alternatively, this estimate implies 0.15 sales per claim, consistent with our aggregate calculation. As this estimate corrects for noncompliance with treatment, aggregating the policy effect using this estimate and a ZIP-code sort based on claims yields a larger aggregate estimate: approximately 242,000 within sample. We focus on the reduced-form estimate for ease of interpretation.

The aggregate estimates are given by 0.76% times the number of tax filers in 2009 with incomes below the phase-out (69 million), and 3.2% times the number of eligible tax filers in 2009 (19 million). These estimates imply that 520,000 or 610,000 households, respectively, were induced by the FTHC to purchase homes. Though using a different source of identification, these estimates are quite close to the range implied by our cross-sectional approach, lending further credibility to each research design.

Method 3: Age distribution approach. We obtain an alternative estimate of the aggregate effect of the FTHC program using variation embedded in the age distribution of first-time buyers, which shifts substantially towards young buyers in 2009. We use this variation to construct our third estimate, which focuses on the total number of young buyers induced by the program.

We first divide the age distribution in 2009 into a young group of those under 35 and a benchmark group of those above 35. We then use the years 2002 through 2007 to estimate the average density in the young group with a simple regression of group shares on age bins. This regression yields a counterfactual predicted density, $f^{\text{young}}_{t\neq 2009}$, and a predicted counterfactual number of young buyers according to the formula, $f^{\text{young}}_{t\neq 2009} = N^{\text{young}}_{2009,cf}/(N^{\text{young}}_{2009,cf} + N^{\text{not young}}_{2009})$. The estimate of induced young buyers is the difference between actual young buyers and counterfactual buyers.

Under the assumption that older buyers are unaffected by the program, this procedure yields an aggregate estimate of 261,000 induced buyers under 35, or approximately 22% of first-time buyers in this age group identified in our data. This estimate provides a lower bound on the overall estimate as the assumption that older buyers are unaffected is too strong. Nevertheless, the age distribution research design provides an alternative source of variation that complements our main design and yields a similar estimate at the group level.

B. *Direct and Indirect Stimulative Effects*

Direct effects. We now provide a back-of-the-envelope calculation of the direct fiscal multiplier using our aggregate estimates of induced home sales. As in Best and Kleven (2017), we abstract from any dynamic effects, and thus our estimates are best interpreted as a lower bound since any post-policy reversal would lower them.[15]

Our baseline estimate of the multiplier uses our in-sample cross-sectional quantity estimates (169K with the conservative counterfactual and 233K with the more aggressive counterfactual). We focus on this sample rather than the aggregate economy because, for this sample, we can construct a quantitative estimate of the indirect effects of the FTHC following Berger et al. (2017). However, for robustness and completeness, we also estimate the overall direct GDP contribution of the FTHC using our extrapolated cross-sectional ag-

[15] We focus on the static effects for three reasons. First, given that we argue that the direct stimulative effects are modest, this assumption is conservative. Second, this assumption has no bearing on our second point, which is that the indirect effects are larger than the direct effects because it affects both proportionately. Third, assuming no reversal is a useful benchmark given that we find little evidence of reversal even two years after the FTHC ended.

gregate estimate and our aggregate RK estimate. The multipliers discussed in this section are summarized in Table VII.

We first consider the income generated for realtors. On average, the realtor fee ranges from 5% to 5.5% of the purchase price,[16] so we use 5.25% in our calculations. When evaluated at the median purchase price for homes in our sample during the policy period ($190,000), this implies that the GDP contribution was $1.7 billion and $2.3 billion using the conservative and aggressive quantity estimates, respectively.[17]

Next, we draw on the literature that estimates the complementary furniture, home improvement, and related expenditures at the time of a new purchase. The literature estimates these expenditures to be approximately 1.9% (Benmelech, Guren, and Melzer (2017)) or 2.8% (Best and Kleven (2017)) of the purchase price. Applying the average of 2.35% to our cross-sectional estimates implies an additional contribution to GDP of $0.8 billion to $1.0 billion using the conservative and aggressive quantity estimates, respectively.[18]

Combining these two calculations, the overall direct contribution to GDP of the FTHC was $2.5 billion in the conservative case and $3.3 billion in the aggressive case, both significantly below the $14 billion cost of versions two and three of the program.[19] These estimates imply that the FTHC generated fiscal impact multipliers of 0.18 in the conservative case and 0.24 in the aggressive case, both much smaller than typical tax multipliers during a recession. Taken together, these calculations lead us to conclude that the direct GDP effects of the program were relatively modest and significantly below its overall cost.[20]

Indirect effects. Section VI presents evidence that the FTHC played an important role in stabilizing housing markets in the aftermath of the Great Recession. One of the main ways in which the policy did so was by increasing home values: a one-standard-deviation increase in program exposure caused the median home to appreciate by approximately $2,400. Given that the FTHC had a significant effect on house prices, natural questions to ask are: first, did the FTHC indirectly stimulate consumption for both claimants and existing homeowners through its effect on house prices, and second, how large are these effects?

[16] Based on data collected by the Department of Justice at https://www.justice.gov/atr/home-prices-and-commissions-over-time (last accessed in August, 2016).

[17] $1.7 billion = 0.0525×190K×169K; $2.3 billion = 0.0525×190K×233K.

[18] $0.80 billion = 0.0235×190K×169K; $1.0 billion = 0.0235×190K×233K.

[19] The total cost of the FTHC program was $20.3 billion and we have 69% of the claims in our sample. The relative concentration in our sample claims versus national home sales accounts for the difference between our in-sample direct multipliers and the full-sample multipliers we compute below using the RKD estimates.

[20] Best and Kleven (2017) find that the stimulative effect of temporary changes in transaction taxes are much larger than ours; their estimates suggest that the impact multiplier is around two. One potential difference is the fact that their policy variation affects all buyers, including repeat buyers who may be more elastic than new homebuyers. Another difference is a relatively large estimate of complementary expenditures around the time of a home purchase.

A recent and growing empirical literature documents large, causal responses of nondurable consumption to house price movements. Using different identification strategies, these studies estimate an elasticity of nondurable consumption to a house price shock (as distinct from housing wealth, which includes leverage[21]) in the range of 0.10 to 0.26.[22] Because house price appreciation affects existing homeowners, it is possible that these indirect effects are large, as housing wealth is the largest component of net worth for most households.[23]

The central theoretical result of Berger et al. (2017) is that despite the many ways in which a change in house prices affects an individual's decision problem, the change in consumption due to an unexpected proportional change in house prices is given by a simple rule-of-thumb formula:

$$\frac{\Delta C_i}{\frac{\Delta P}{P}} = MPC_i \cdot (PH_{i-1}(1-\delta)), \tag{IA2}$$

where MPC_i is the individual marginal propensity to consume (MPC) out of transitory income shocks and $PH_{i-1}(1-\delta)$ is the value of the individual's home after depreciation. Given estimates of these objects, we can aggregate them across households and areas and evaluate the size of the indirect effects under a variety of assumptions.

We proceed as above by choosing the bottom 1% of ZIPs as the counterfactual group and computing an aggregate house price effect for other groups relative to this group. Recall that the standardized exposure is 0.85 for the bottom centile and increases to 7.58 for the highest one. Thus, for each exposure group g, the aggregate percentage change in prices for that group induced by the program is

$$\frac{\Delta P_g}{P_g} = \beta_p \times (e_g - e_{g,low}), \tag{IA3}$$

[21] This is an important distinction. A shock to housing net worth is typically much larger than a shock to house prices because of leverage. To be more specific, a 100% increase in price is a 500% increase in housing wealth if leveraged 80%. This means that the housing wealth elasticity is significantly greater than the elasticity of consumption to a change in house prices. In Berger et al. (2017) we explicitly scale down the estimates in Mian, Rao, and Sufi (2013), who measure the elasticity of consumption to a housing wealth shock, so that it is comparable to the elasticity of consumption to a change in house prices, the object of interest in Berger et al. (2017). This correction lowers the elasticity estimate in Mian, Rao, and Sufi (2013) from 0.5-0.8 to 0.13-0.26. For more details, see footnote 12 in Berger et al. (2017).

[22] See, for example, Mian, Rao, and Sufi (2013), Stroebel and Vavra (2019), and Kaplan, Mitman, and Violante (2016).

[23] A separate question is whether the credit itself affected the consumption of claimants. We focus on the response of consumption to an equilibrium change in house prices rather than this direct effect of receiving the credit because (a) the former effect is likely larger and (b) the direct effect is ambiguous. If the credit does not affect the size of transacted home, the direct consumption effect could be large due to the cash transfer to households. If buyers adjust their home purchase size, direct consumption effects would be smaller.

where β_p is the coefficient (0.01077) from the long-difference price regression in Table VI. We apply this price growth factor to the average median house price in each group and accumulate over all owner-occupied housing units in each group under an assumed value for the marginal propensity to consume. We then apply equation (IA2) to infer the aggregate change in consumption induced by the policy. These calculations are rough but informative for the magnitude of potential indirect effects.

To derive a ballpark estimate of this effect, we apply the rule-of-thumb formula in Berger et al. (2017) to our empirical setting. We use our within-sample cross-sectional estimates because we are unable to implement the formula in equation (IA2) for the entire sample. We conservatively assume that the MPC out of housing wealth is 0.05 for households with LTVs below 80%, as they have enough equity to borrow against even in tight credit markets, and zero for households with little or negative equity LTV.[24] This reduces to scaling equation (IA2) by the fraction of households with LTV below 80% in a given group g using data from January 2010. To be conservative, we further assume that only house prices in the starter-home market (one- to three-bedroom homes) are affected by the policy. Taken together, these assumptions imply that aggregate consumption would have increased by \$4.2 billion with a conservative baseline and \$6.0 billion with a more aggressive baseline. These estimates apply within our estimation sample and only include the indirect effects of the FTHC on the starter-home market. Assuming instead that the entire housing stock is affected by the policy, the effect is \$7.3 billion in the conservative case and \$10.3 billion in the aggressive case.

This exercise is too coarse to permit emphasis of a particular number; the key takeaway is that these effects can be as large or possibly much larger than the program's direct stimulative effects. Even our conservative estimate implies a multiplier for the indirect effect of 0.30 in the conservative case and 0.43 in the aggressive case, which are considerably larger than the direct effect. Were we to apply the indirect effect to the entire housing stock, these multipliers would rise to 0.53 and 0.73, respectively. Combining our preferred direct and indirect effects yields a total fiscal impact multiplier of the FTHC of 0.48 in the conservative case and 0.67 in the aggressive case, most of which comes from the indirect effect—the component typically ignored in most policy evaluations.

Robustness. It is possible to compute an estimate of the overall direct GDP contribution that uses both our full-sample cross-sectional aggregate estimate and our aggregate RK estimate because this calculation only requires having an estimate of the aggregate quantity effect of the policy. Using our full-sample cross-sectional aggregate estimate, we find that the direct contribution to GDP is \$5.9 billion in the conservative case and \$8.2 billion in the aggressive case.[25] Using our RK aggregate estimate, we find that the direct contribution added \$7.5 billion using the estimate from the population of tax filers and \$8.8 billion using

[24]The average MPC out of housing wealth of the nine low-LTV empirical studies discussed in Ganong and Noel (2018) (see Figure 3) is 7.6 cents. The average MPC out of housing wealth of the studies discussed in Carroll et al. (2011) is 5.5 cents.

[25]\$5.9 billion = $(0.0525 + 0.0235)\times 190\text{K}\times 412\text{K}$; \$8.2 billion = $(0.0525 + 0.0235)\times 190\text{K}\times 568\text{K}$.

the estimate from the sample of eligibles only.[26] Since the overall cost of the FTHC program was \$20.3 billion, these dollar amounts imply that the direct impact multiplier of the FTHC program is 0.29 in the conservative case and 0.40 in the aggressive case using our full-sample cross-sectional estimates. Using the aggregate RK estimate, the direct impact multiplier is 0.37 for the full population sample and 0.43 for the eligibles-only sample.

[26] \$7.5 billion $= (0.0525 + 0.0235) \times 190\text{K} \times 520\text{K}$; \$8.8 billion $= (0.0525 + 0.0235) \times 190\text{K} \times 610\text{K}$.

REFERENCES

Benmelech, Efraim, Adam Guren, and Brian T. Melzer, 2017, Making the house a home: The stimulative effect of home purchases on consumption and investment, NBER Working Paper No. 23570.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra, 2017, House prices and consumer spending, *Review of Economic Studies* 85, 1502–1542.

Best, Michael Carlos, and Henrik Jacobsen Kleven, 2017, Housing market responses to transaction taxes: Evidence from notches and stimulus in the uk, *Review of Economic Studies* 85, 157–193.

Bogin, Alexander N., William M. Doerner, and William D. Larson, 2016, Local house price dynamics: New indices and stylized facts, FHFA Working Paper No. 16-01.

Carroll, Christopher D., Misuzu Otsuka, and Jiri Slacalek, 2011, How large are housing and financial wealth effects? A new approach, *Journal of Money, Credit and Banking* 43, 55–79.

Department of Housing and Urban Development, 2011, Annual Report to Congress Regarding the Financial Status of the FHA Mutual Mortgage Insurance Fund Fiscal Year 2011, http://portal.hud.gov/hudportal/documents/huddoc?id=FHAMMIFundAnnRptFY11No2.pdf.

Engelhardt, Gary, 1996, Consumption, down payments, and liquidity constraints, *Journal of Money, Credit and Banking* 28, 255–271.

Fuster, Andreas, and Basit Zafar, 2016, To buy or not to buy: Consumer constraints in the housing market, *American Economic Review: Papers and Proceedings* 106, 636–40.

Ganong, Peter, and Pascal Noel, 2018, Liquidity vs. wealth in household debt obligations: Evidence from housing policy in the Great Recession, NBER Working Paper No. 24964.

Goodman, Laurie, Ellen Seidman, and Jun Zhu, 2014, FHA loan limits: What areas are the most affected?, Urban Institute.

Kaplan, Greg, Kurt Mitman, and Gianluca Violante, 2016, Non-durable consumption and housing net worth in the Great Recession: Evidence from easily accessible data, NBER Working Papers 22232.

Mian, Atif, Kamelesh Rao, and Amir Sufi, 2013, Household balance sheets, consumption, and the economic slump, *Quarterly Journal of Economics* 128, 1687–1726.

Mian, Atif, and Amir Sufi, 2009, The consequences of mortgage credit expansion: Evidence from the U.S. mortgage default crisis, *Quarterly Journal of Economics* 124, 1449–1496.

Mian, Atif, and Amir Sufi, 2012, The effects of fiscal stimulus: Evidence from the 2009 Cash for Clunkers program, *Quarterly Journal of Economics* 127, 1107–1142.

Stein, Jeremy, 1995, Prices and trading volume in the housing market: A model with down-payment effects, *Quarterly Journal of Economics* 110, 379–406.

Stroebel, Johannes, and Joseph Vavra, 2019, House prices, local demand, and retail prices, *Journal of Political Economy* 127, 1391–1436.

IV. Additional Figures and Tables

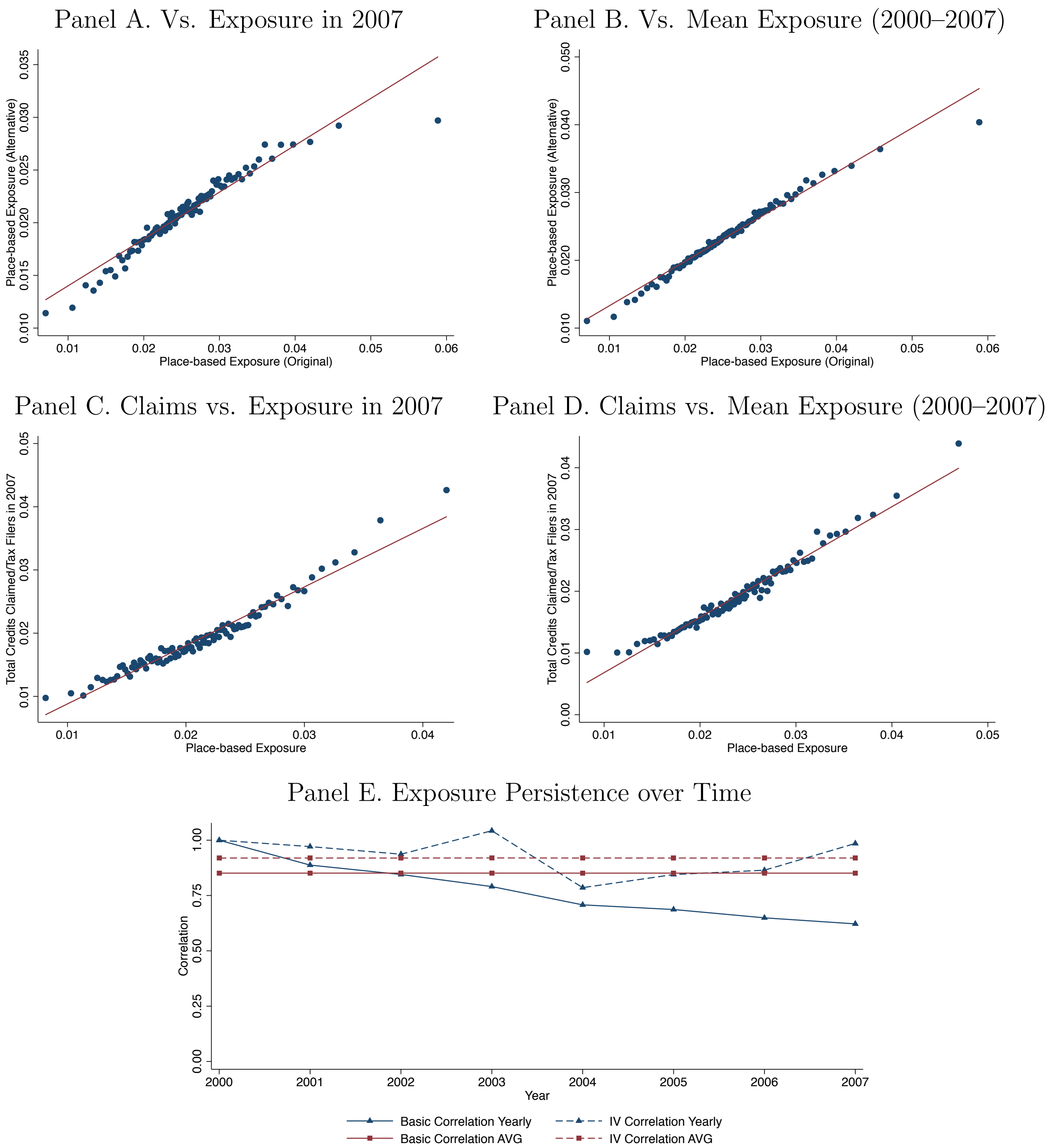


Figure IA.1. FTHC exposure is persistent over time. The panels in this figure demonstrate the persistence of FTHC exposure at the ZIP level over time. We construct two alternative measures of exposure. The first uses the number of first-time homebuyers in 2007 scaled by the total number of filing units. The second uses the population-weighted average of exposure over the period 2000 to 2007. Panels A and B plot binscatters of these measures against our baseline exposure measure. Panels C and D plot binscatters of FTHC claims versus these alternative measures. Panel E plots the raw correlation of these measures with baseline exposure over time and a version that instruments for the alternative exposure measure with the total number of buyers scaled by population in 2000 to remove measurement error. Correlations are weighted by total home sales in 2007.

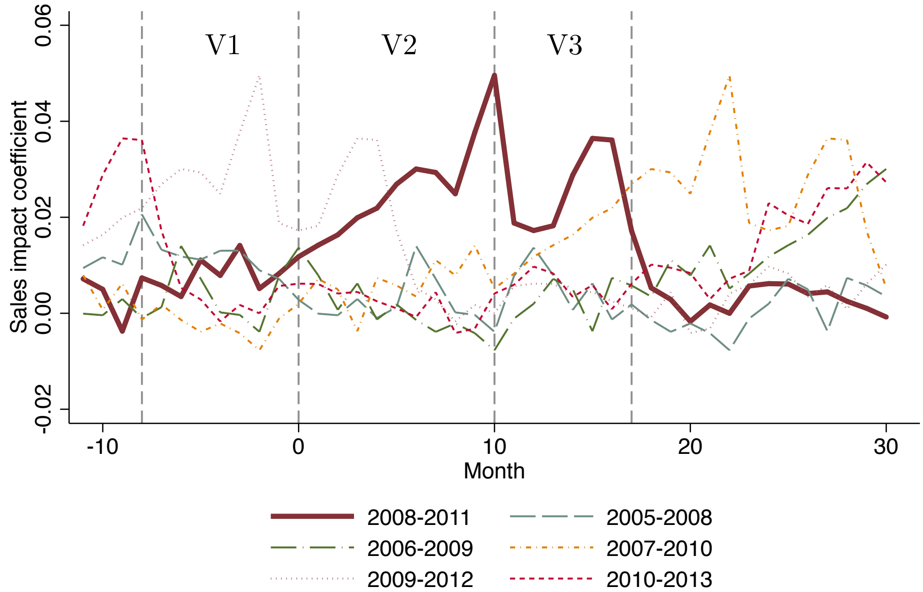


Figure IA.2. Placebo coefficients. This figure depicts a placebo test for whether seasonality accounts for the spikes in the home sales distribution. The test estimates month-by-month regressions and plots coefficients from the control specification in Figure 5, Panel B, emphasized with a bold line, along with equivalent regressions shifted backward in time to start in 2005, 2006, and 2007, and shifted forward to start in 2009 and 2010. The vertical markers correspond to the FTHC loan program (V1), the start of the FTHC grant program (V2), the scheduled expiration of the FTHC grant program, and the actual expiration of the FTHC grant program (V3).

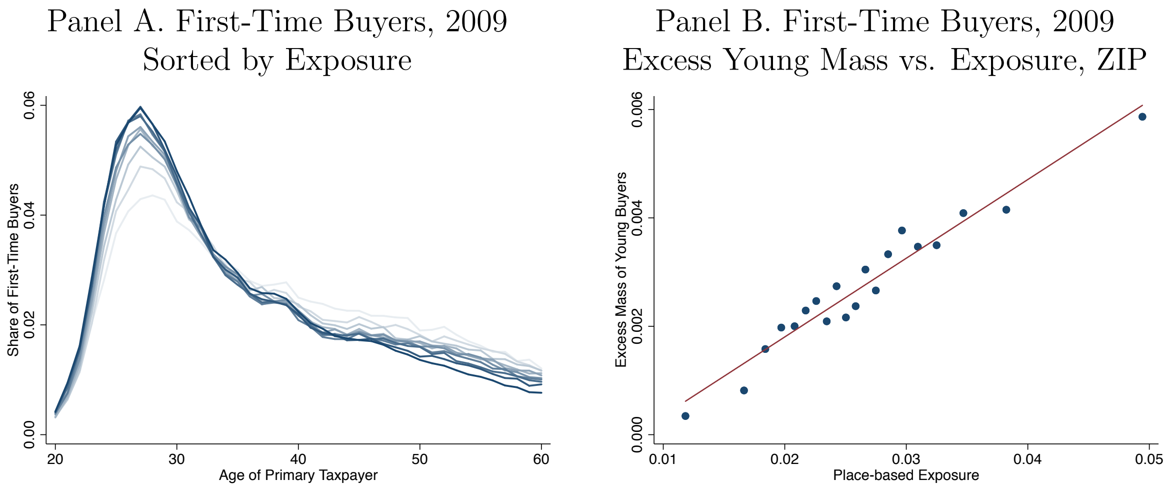


Figure IA.3. Policy shift in the age distribution of first-time buyers (appendix). Panel A plots the age distributions of first-time homebuyers in 2009 identified using information returns across exposure deciles. For each line, the denominator is the total number of first-time homebuyers within an exposure decile. Darker colors indicate higher exposure deciles. Panel B shows the correlation between the shift in the age distribution in 2009 and program exposure. We decompose the national shift in the age distribution in 2009 into contributions from each ZIP. For each ZIP, we compute the difference between the ratio of buyers aged 30 or younger to total new homebuyers in 2009 versus the average ratio of buyers aged 30 or younger to total new homebuyers in other years. We then plot binned bivariate sums (20 bins) of these ZIP-level contributions against average exposure in each bin.

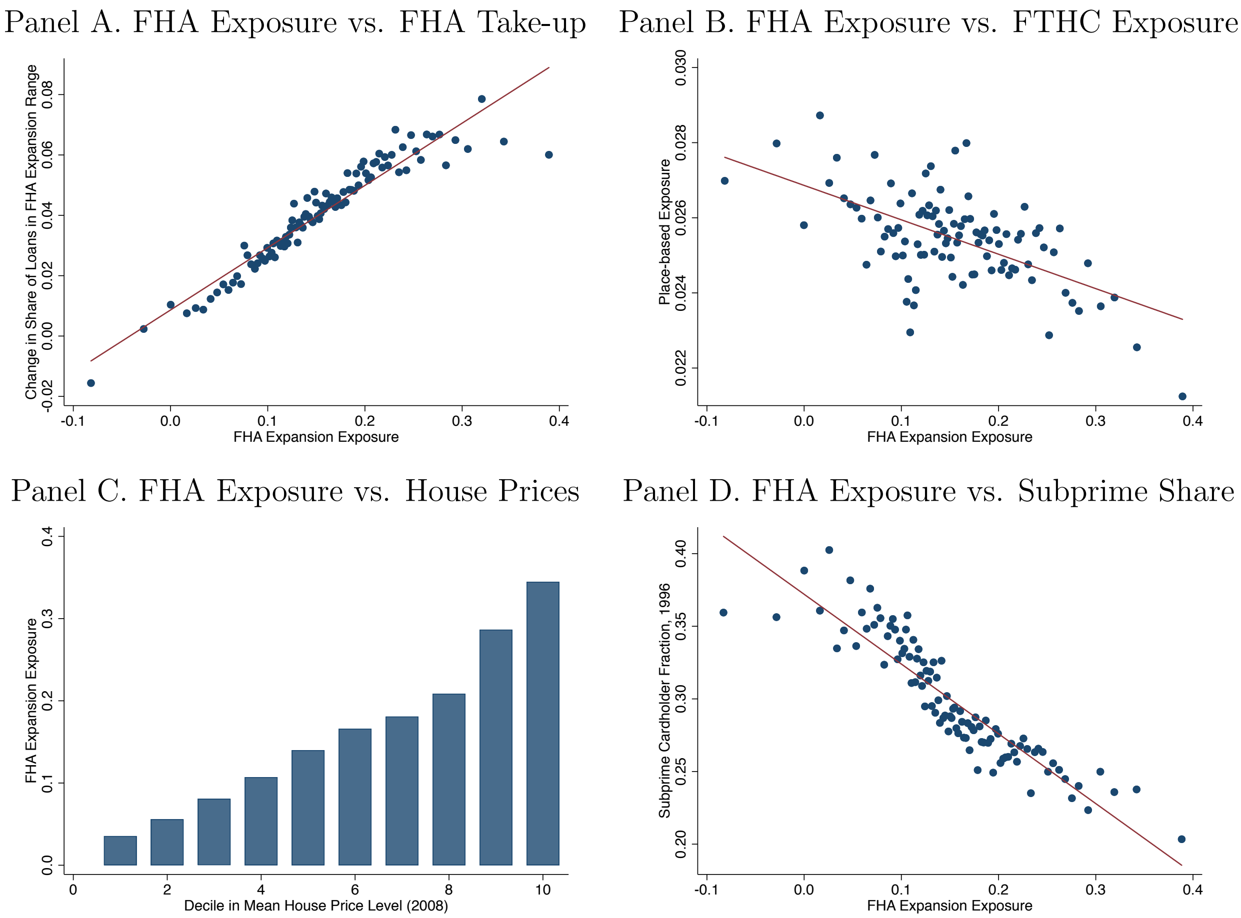


Figure IA.4. FHA exposure is correlated with FHA take-up, not FTHC exposure. The panels in this figure show that the FHA expansion in 2009 targeted different ZIPs than the FTHC program, as proxied by our exposure measure. We measure FHA expansion exposure as the average share of loans in 2004 to 2007 in HMDA in the range between the pre-expansion FHA limit and the post-expansion FHA limit. Panel A plots a binscatter (100 bins) between FHA exposure at the ZIP level and the change in the share of FHA originations in HMDA in the FHA expansion range, confirming that our measure is a good predictor of the FHA expansion. Panel B plots a binscatter (100 bins) of FHA exposure versus FTHC exposure, conditional on CBSA fixed effects, revealing a weak negative correlation. Panel C shows that this negative correlation is explained in part by the fact that FHA exposure is highest in high-house-price areas, which tend to have fewer first-time homebuyers. Consistent with this fact, Panel D plots a binscatter (100 bins) of FHA exposure versus the subprime share in 1996, suggesting that the FHA expansion also targeted areas that had relatively fewer subprime borrowers.

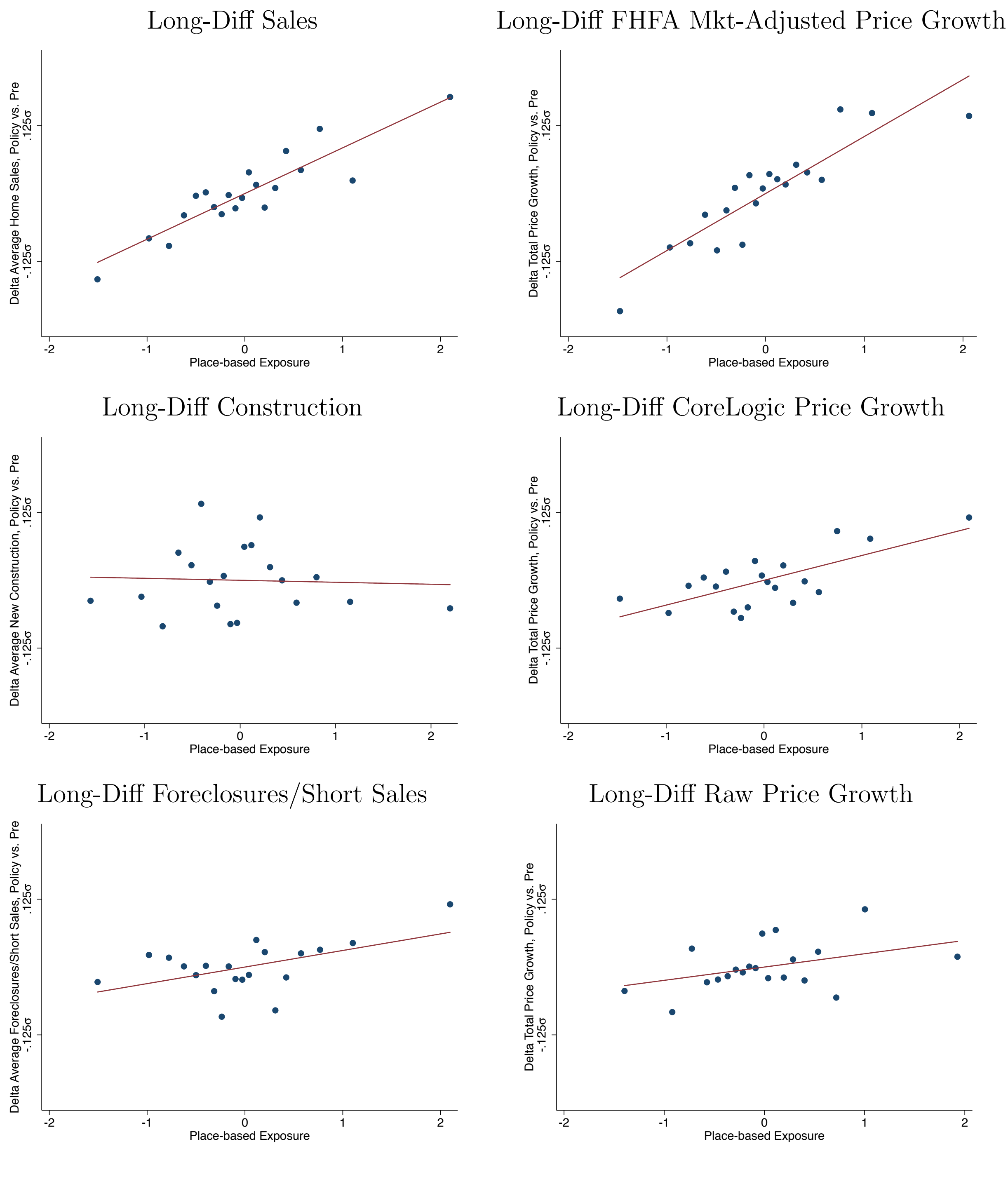
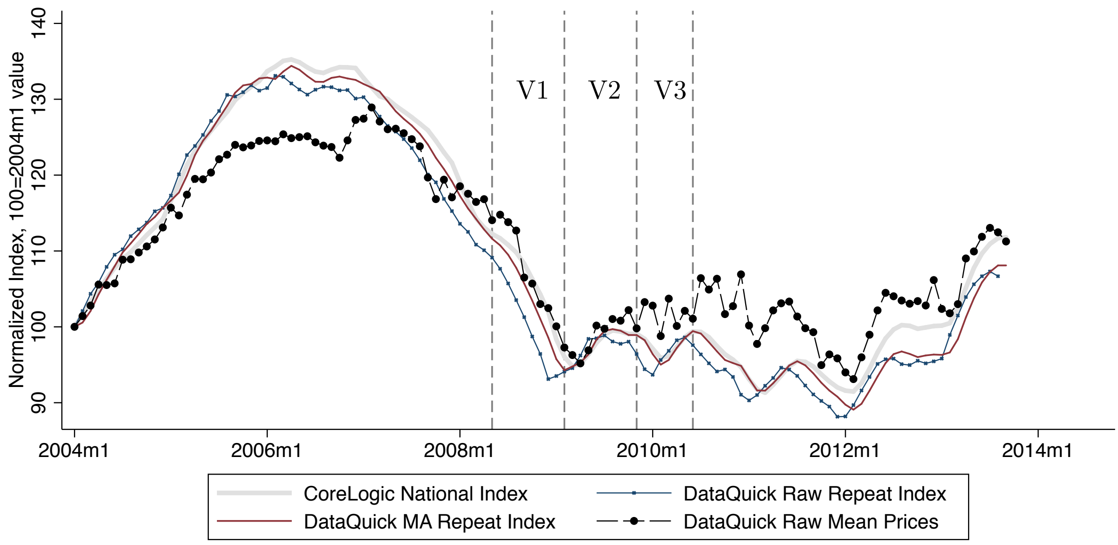
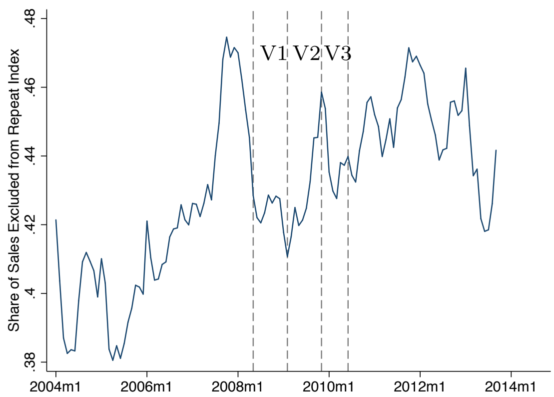


Figure IA.5. Long-difference binscatters. The panels in this figure plot binscatters of long-difference effects of the FTHC on various outcomes from Tables IV and VI versus program exposure. Exposure is defined as the number of first-time homebuyers in an area in 2000. The y-axis is scaled in terms of standard deviations of the left-hand-side variable. Exposure is normalized by its cross-sectional standard deviation.

Panel A. Comparing Price Indices and Raw Prices



Panel B. Transactions Excluded from Index



Panel C. House Vintages, Excluded vs. Included

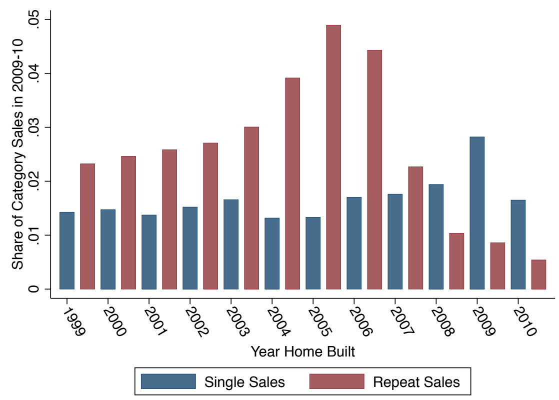


Figure IA.6. Aggregate prices and price index composition. The panels in this figure explore the composition of repeat-sales price indices during the time around the FTHC. Panel A plots the CoreLogic national house price index and three separate price indexes that we build from DataQuick transaction data. The CoreLogic index is based on repeat-sales pairs and incorporates a three-month moving average to smooth the aggregate data. The raw repeat index follows published methodology from CoreLogic and S&P Case-Shiller for including sales pairs and estimating price indices. The MA repeat index applies a three-month moving average to the raw repeat index. The raw mean price index includes all transactions, winsorized at the top and bottom 1% to reduce the influence of outliers. The vertical markers correspond to the FTHC loan program (V1), the start of the FTHC grant program (V2), the scheduled expiration of the FTHC grant program, and the actual expiration of the FTHC grant program (V3). Panel B reports the share of transactions by month that are excluded from the raw repeat index because those properties transact only once during the estimation window. Panel C plots the density of house vintages for excluded (single) and included (repeat) transactions using information from assessor files.

Table IA.I

The Effect of the FTHC on Home Sales (Panel Specification)

This table presents the average monthly effects of the FTHC on home sales for ZIPs pooled over different policy windows with various specifications. We run panel regressions, weighted by average monthly home sales in 2007, of the form

$$y_{it} = \alpha_i + \delta_{CBSA,t} + \sum_P \beta_P Exposure_i + \sum_t \gamma_t X_i + \varepsilon_{it},$$

where y_i is the log of monthly home sales in area i between 2004 and 2013m5. In controls specifications, X_i is a set of controls that include log population, the average unemployment rate from 2006 through 2010, the log of average gross income, the subprime share in 1996, the average share of subprime originations from 2004 to 2007, exposure to the FHA expansion, and exposure to the HARP and HAMP programs. Exposure is normalized by its cross-sectional standard deviation. All columns include CBSA-by-month fixed effects. Column (2) presents unweighted regressions. Column (3) excludes Arizona, California, and Nevada. Column (4) restricts the sample to areas with average home sales in 2007 above the 10th percentile. Standard errors are clustered at the ZIP level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively.

	Base (1)	No wgts (2)	Ex sand (3)	$\overline{\text{Sales}}$ > P10 (4)
Pre-policy	-0.003	-0.012**	-0.002	-0.003
2006m4-2008m4	(0.004)	(0.003)	(0.004)	(0.004)
FTHC Loan Program	0.008*	-0.001	0.009*	0.008
2008m5-2008m12	(0.005)	(0.004)	(0.005)	(0.005)
FTHC Grant Program	0.024***	0.018***	0.026***	0.025***
2009m1-2010m5	(0.005)	(0.004)	(0.005)	(0.005)
Reversal	-0.007	-0.012***	-0.006	-0.007
2010m6-2011m5	(0.005)	(0.005)	(0.006)	(0.006)
Post-reversal	-0.0	-0.006	0.003	0.0
2011m6-2013m5	(0.005)	(0.005)	(0.006)	(0.005)
Observations	959916	959916	816013	868382
R^2	0.852	0.828	0.854	0.847
Controls	Yes	Yes	Yes	Yes
CBSA-Month and ZIP FE	Yes	Yes	Yes	Yes

Table IA.II

Total Number of Observations in Data Set through Each Filter

The number of unique geographic units in the data set (ZIPs, counties, or CBSAs) are in parentheses.

	ZIPs	Counties	CBSAs
Geo-month observations (transaction counts in parentheses)			
Matched between assessor and transaction data	2,716,338	150,859	55,348
	(124.4 M)	(124.4 M)	(117.5 M)
+ Arm's length transactions w/valid geo, month	2,540,700	145,776	53,575
	(70.51 M)	(70.51 M)	(67.26 M)
+ Cleaned resales & new sales over 2004 to 2013	1,423,144	85,724	29,509
	(37.28 M)	(37.28 M)	(35.37 M)
+ Nondistress resales	1,371,576	85,560	29,481
	(21.68 M)	(21.68 M)	(20.37 M)
+ Time series 90%+ complete over 2006 to 2013	1,042,080	75,687	27,107
	(20.41 M)	(21.08 M)	(20.05 M)
+ Matched to exposure variables and covariates	1,019,086	75,570	26,756
	(19.95 M)	(21.07 M)	(19.81 M)
Unique geographic units in dataset			
Matched between assessor and transaction data	19,240	973	295
+ Arm's length transactions w/valid geo, month	18,351	970	295
+ Cleaned resales & new sales over 2004 to 2013	17,671	941	295
+ Nondistress resales	17453	941	295
+ Time series 90%+ complete over 2006 to 2013	9,082	664	235
+ Matched to exposure variables and covariates	8,883	663	232

Table IA.III

The Effect of the FTHC on Starter Homes vs. Large Homes

This table presents regressions of the same form as those in Table III. We divide the home sales series into "starter" homes—defined as those with one, two, or three bedrooms—and large homes—defined as those with four or more bedrooms. We run the ZIP-level specifications separately for each series. The analysis sample here is the subset of the main analysis sample where fewer than 5% of transactions between 2004 and 2013 have missing bedrooms data. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively.

Panel A. 1-3 Bedrooms, ZIP			Panel B. 4+ Bedrooms, ZIP		
	No Controls (1)	CBSA FE (2)		No Controls (1)	CBSA FE (2)
Pre-policy	0.008	0.012*	Pre-policy	-0.008	-0.005
2007m9-2009m1	(0.008)	(0.007)	2007m9-2009m1	(0.006)	(0.007)
Observations	2971	2904	Observations	2120	2063
R^2	0.003	0.458	R^2	0.004	0.455
Policy	0.013	0.023***	Policy	-0.002	0.002
2009m2-2010m6	(0.011)	(0.006)	2009m2-2010m6	(0.008)	(0.008)
Observations	2971	2904	Observations	2120	2063
R^2	0.004	0.516	R^2	0.000	0.455
Post-policy	0.005	0.011*	Post-policy	-0.007	-0.007
2010m7-2011m11	(0.012)	(0.006)	2010m7-2011m11	(0.008)	(0.008)
Observations	2971	2904	Observations	2120	2063
R^2	0.000	0.593	R^2	0.002	0.495
Early policy	0.004	0.018***	Early policy	-0.004	0.001
2009m2-2009m9	(0.009)	(0.005)	2009m2-2009m9	(0.007)	(0.007)
Observations	2971	2904	Observations	2120	2063
R^2	0.001	0.467	R^2	0.001	0.430
Spike 1	0.029*	0.036***	Spike 1	0.003	0.005
2009m10-2009m12	(0.014)	(0.010)	2009m10-2009m12	(0.008)	(0.009)
Observations	2970	2903	Observations	2120	2063
R^2	0.011	0.473	R^2	0.000	0.427
Spike 2	0.018	0.025***	Spike 2	-0.001	-0.000
2010m4-2010m6	(0.012)	(0.007)	2010m4-2010m6	(0.008)	(0.009)
Observations	2968	2901	Observations	2119	2062
R^2	0.005	0.433	R^2	0.000	0.401
Controls	No	Yes	Controls	No	Yes
CBSA FE	No	Yes	CBSA FE	No	Yes

Table IA.IV

Summary Statistics for Regression Kink Analysis in Table V

This table presents summary statistics and balance statistics for the regression kink analysis sample.

	Full Sample	Below Kink	Above Kink
Age	44.50	44.32	44.71
	[9.69]	[9.81]	[9.54]
Married	0.60	0.58	0.62
	[0.49]	[0.49]	[0.49]
First-time Home Buyer	0.0143	0.0154	0.0130
	[0.12]	[0.12]	[0.11]
First-time Home Buyer Credit ($)	69.55	81.39	55.09
	[723.10]	[797.83]	[619.44]
First-time Home Buyer Credit (> 0) ($)	7370.76	7763.06	6754.93
	[1264.02]	[1038.10]	[1339.20]
Existing Owner	0.75	0.74	0.77
	[0.43]	[0.44]	[0.42]
Adjusted Gross Income ($2013)	129639.90	124467.80	135955.20
	[40434.93]	[40346.72]	[39637.60]
Wages ($2013)	115564.80	111291.80	120782.30
	[47065.82]	[46107.75]	[47693.56]
Family Size	3.14	3.17	3.12
	[1.54]	[1.55]	[1.53]
Distance to Kink ($ 00s)	-5.33	-41.34	38.65
	[46.01]	[23.11]	[23.08]
Observations	3858042	2121006	1737036

Table IA.V

The Effect of the FTHC on Home Sales (Alternative Exposure, 2007)

This table presents regressions of the average monthly effects of the FTHC on different categories of home sales, where we replace baseline exposure with exposure measured in 2007. We run cross-sectional regressions, weighted by average monthly home sales in 2007, of the form

$$y_i = \alpha_{CBSA} + \beta Exposure'_i + \gamma X_i + \varepsilon_i,$$

where y_i is a housing market outcome in area i over the relevant time period. In the first row, the outcome is the difference in average monthly nondistress home resales for the policy period versus the 17-month pre-period. In the second row, the outcome is the difference in average monthly new construction sales for the policy period versus the pre-period. In the third row, the outcome is the difference in average monthly foreclosures and short sales for the policy period versus the pre-period. Exposure is normalized by its cross-sectional standard deviation. All columns include CBSA fixed effects and controls that include log population, the average unemployment rate from 2006 through 2010, the log of average gross income, the subprime share in 1996, the average share of subprime originations from 2004 to 2007, exposure to the FHA expansion, and exposure to the HARP and HAMP programs. Column (2) presents unweighted regressions. Column (3) excludes Arizona, California, Florida, and Nevada. Column (4) trims the left-hand-side variable at the 5th and 95th percentiles. Column (5) restricts the sample to areas with average home sales in 2007 above the 10th percentile. Columns (6) and (7) divide the sample of ZIPs into the bottom three ("Low p") and top three ("High p") deciles in median house prices during 2008. Standard errors are clustered at the CBSA level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively.

	CBSA FE (1)	No wgts (2)	Ex sand (3)	Trimmed (4)	$\overline{\text{Sales}}$ > P10 (5)	Low p (6)	High p (7)
			LHS is Long-Diff Sales				
Coefficient	0.021***	0.023***	0.019***	0.014***	0.021***	0.025***	-0.006
	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)	(0.007)
Observations	8689	8689	6663	7832	7824	2432	2372
R^2	0.445	0.463	0.361	0.445	0.451	0.626	0.434
			LHS is Long-Diff Construction				
Coefficient	-0.009	-0.009	-0.008	0.005	-0.009	-0.003	-0.009
	(0.007)	(0.006)	(0.009)	(0.005)	(0.008)	(0.021)	(0.028)
Observations	4721	4721	3411	4243	4522	1150	1156
R^2	0.121	0.117	0.128	0.161	0.123	0.208	0.160
			LHS is Long-Diff Foreclosures & Short Sales				
Coefficient	0.039*	0.057**	0.032	0.021	0.039*	0.041	-0.011
	(0.022)	(0.022)	(0.025)	(0.017)	(0.023)	(0.029)	(0.044)
Observations	8504	8504	6478	7663	7704	2393	2313
R^2	0.343	0.320	0.232	0.433	0.349	0.433	0.272
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table IA.VI

The Effect of the FTHC on Home Sales (Alternative Exposure, 2000–2007 Mean)

This table presents regressions of the average monthly effects of the FTHC on different categories of home sales, where we replace baseline exposure with exposure measured as mean exposure in 2000 to 2007. We run cross-sectional regressions, weighted by average monthly home sales in 2007, of the form

$$y_i = \alpha_{CBSA} + \beta Exposure'_i + \gamma X_i + \varepsilon_i,$$

where y_i is a housing market outcome in area i over the relevant time period. In the first row, the outcome is the difference in average monthly nondistress home resales for the policy period versus the 17-month pre-period. In the second row, the outcome is the difference in average monthly new construction sales for the policy period versus the pre-period. In the third row, the outcome is the difference in average monthly foreclosures and short sales for the policy period versus the pre-period. Exposure is normalized by its cross-sectional standard deviation. All columns include CBSA fixed effects and controls that include log population, the average unemployment rate from 2006 through 2010, the log of average gross income, the subprime share in 1996, the average share of subprime originations from 2004 to 2007, exposure to the FHA expansion, and exposure to the HARP and HAMP programs. Column (2) presents unweighted regressions. Column (3) excludes Arizona, California, Florida, and Nevada. Column (4) trims the left-hand-side variable at the 5th and 95th percentiles. Column (5) restricts the sample to areas with average home sales in 2007 above the 10th percentile. Columns (6) and (7) divide the sample of ZIPs into the bottom three ("Low p") and top three ("High p") deciles in median house prices during 2008. Standard errors are clustered at the CBSA level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively.

	CBSA FE (1)	No wgts (2)	Ex sand (3)	Trimmed (4)	$\overline{\text{Sales}}$ > P10 (5)	Low p (6)	High p (7)
				LHS is Long-Diff Sales			
Coefficient	0.030***	0.037***	0.023***	0.018***	0.030***	0.033***	-0.000
	(0.004)	(0.006)	(0.003)	(0.003)	(0.004)	(0.005)	(0.006)
Observations	8692	8692	6664	7834	7825	2432	2375
R^2	0.454	0.472	0.366	0.449	0.459	0.632	0.434
				LHS is Long-Diff Construction			
Coefficient	0.000	0.007	0.002	0.013**	-0.000	0.008	0.010
	(0.007)	(0.007)	(0.009)	(0.006)	(0.008)	(0.021)	(0.021)
Observations	4721	4721	3411	4243	4522	1150	1156
R^2	0.121	0.117	0.127	0.162	0.122	0.208	0.160
				LHS is Long-Diff Foreclosures & Short Sales			
Coefficient	0.039	0.043*	0.038	0.031*	0.039	0.051	-0.026
	(0.024)	(0.024)	(0.027)	(0.016)	(0.025)	(0.032)	(0.039)
Observations	8507	8507	6479	7664	7705	2393	2316
R^2	0.343	0.319	0.232	0.433	0.349	0.433	0.272
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table IA.VII

The Effect of the FTHC on Home Sales (Reversal Period)

This table presents regressions of the average monthly effects of the FTHC on home sales during the reversal period, defined as the 12 months after the expiration of program. We run cross-sectional regressions, weighted by average monthly home sales in 2007, of the form

$$y_i = \alpha_{CBSA} + \beta Exposure_i + \gamma X_i + \varepsilon_i,$$

where y_i is the difference in average monthly nondistress home resales for the reversal period versus the 17-month pre-period. Exposure is normalized by its cross-sectional standard deviation. All columns include CBSA fixed effects and controls that include log population, the average unemployment rate from 2006 through 2010, the log of average gross income, the subprime share in 1996, the average share of subprime originations from 2004 to 2007, exposure to the FHA expansion, and exposure to the HARP and HAMP programs. Columns (2) and (3) divide the sample of ZIPs into the bottom three ("Low p") and top three ("High p") deciles in median house prices during 2008. Columns (4) and (5) divide the sample of ZIPs into the bottom three ("Low Start") and top three ("High Start") deciles in the starter-home (defined as in Table IA.III) share of sales in sample years prior to 2009. Columns (6) and (7) divide the sample of ZIPs into the bottom three ("Low Hold") and top three ("High Hold") deciles in the average pre-policy holding time for first-time homebuyers, identified in the Federal Reserve Bank of New York Consumer Credit Panel, constructed from Equifax. Standard errors are clustered at the CBSA level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively.

	CBSA FE (1)	Low p (2)	High p (3)	Low Start (4)	High Start (5)	Low Hold (6)	High Hold (7)
	LHS is Long-Diff Post						
Coefficient	-0.000	0.003	-0.010	-0.003	-0.006	-0.008	0.004
	(0.004)	(0.005)	(0.011)	(0.013)	(0.009)	(0.006)	(0.005)
Equality test		p = 0.251		p = 0.834		p = 0.067	
Observations	8692	2431	2375	833	850	2948	2804
R^2	0.533	0.623	0.415	0.631	0.713	0.551	0.546
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table IA.VIII

The Effect of the FTHC on House Prices (Appendix)

This table presents regressions of the cumulative effects of the FTHC on different measures of house price growth. We run cross-sectional regressions, weighted by average monthly home sales in 2007, of the form

$$y_i = \alpha_{CBSA} + \beta Exposure_i + \gamma X_i + \varepsilon_i,$$

where y_i is a housing market outcome in area i over the relevant time period. In the first row, the outcome is cumulative annual log price differences during the policy period minus differences during the 17-month pre-period from a hedonic price index based on our DataQuick data. Within our sample of ZIPs, we estimate a hedonic price model where the (5% winsorized) log value of the transaction price is the left-hand-side variable with deciles of lot size, ventiles of square footage, ventiles of age, indicators for the number of bedrooms and bathrooms (rounded to the nearest integer), and month-by-ZIP fixed effects included as covariates. The exponentiated values of the month-by-ZIP fixed effects are the house price index. We winsorize the indices at the 5% level. In the second row, the outcome is cumulative monthly log price differences from DataQuick mean prices, winsorized at the 5% level, for the same pre- and policy periods. In all cases, we multiply the left-hand side by 100 so the treatment effect units are percentage points of growth per standard deviation change in program exposure. Exposure is normalized by its cross-sectional standard deviation. Column (2) presents unweighted regressions. Column (3) excludes Arizona, California, Florida, and Nevada. Column (4) trims the left-hand-side variable at the 5th and 95th percentiles. Column (5) restricts the sample to areas with average home sales in 2007 above the 10th percentile. Columns (6) and (7) divide the sample of ZIPs into the bottom three ("Low p") and top three ("High p") deciles in 2008 median house prices. Standard errors are clustered at the CBSA level. *, **, and *** denote significance at 0.10, 0.05, and 0.01 levels, respectively.

	CBSA FE (1)	No wgts (2)	Ex sand (3)	Trimmed (4)	$\overline{\text{Sales}}$ > P10 (5)	Low p (6)	High p (7)
	LHS is Long-Diff Price Growth (Hedonic)						
Coefficient	1.927***	2.29***	1.699***	2.31***	1.849***	3.061*	1.487
	(0.584)	(0.573)	(0.63)	(0.482)	(0.597)	(1.615)	(1.008)
Observations	5510	5510	4462	4955	5021	1537	1561
R^2	0.212	0.223	0.163	0.227	0.214	0.227	0.362
	LHS is Long-Diff Raw Price Growth (DataQuick)						
Coefficient	1.595	1.785	0.609	0.117	1.549	0.575	0.775
	(1.061)	(1.222)	(1.285)	(0.82)	(1.085)	(4.01)	(1.677)
Observations	7308	7308	5372	6573	6851	2071	2198
R^2	0.088	0.092	0.078	0.1	0.089	0.147	0.147
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CBSA FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes